La radiologia medica (2018) 123:515–523
https://doi.org/10.1007/s11547-018-0872-9

COMPUTED TOMOGRAPHY

Comparison of ultrasonography and computed tomography for diagnosing diffuse thyroid disease: a multicenter study

Dong Wook Kim[1] · Yoo Jin Lee[1] · Hye Shin Ahn[2] · Hye Jin Baek[3] · Ji Hwa Ryu[4] · Taewoo Kang[5]

Received: 9 August 2017 / Accepted: 6 March 2018 / Published online: 10 March 2018

Abstract

Purpose To compare the diagnostic performance of ultrasonography (US) and computed tomography (CT) for diagnosing incidentally detected diffuse thyroid disease (DTD) in patients who underwent thyroid surgery using multicenter data.

Methods Between July and December 2016, a total of 177 patients who underwent preoperative thyroid US and neck CT, and subsequent thyroid surgery at 4 participating institutions, were reviewed. US and CT images in each case were retrospectively reviewed by a radiologist at each institution, and classified into one of the following four categories based on US and CT features: no DTD; indeterminate; suspicious for DTD; and DTD. The diagnostic accuracy of US and CT were calculated at each institution by comparison with histopathological results.

Results Respective US and CT classifications in the 177 patients were no DTD in 75 and 71, indeterminate in 46 and 34, suspicious for DTD in 28 and 31, and DTD in 28 and 41. Among the histopathological results, 113 patients had normal thyroid parenchyma, 23 had Hashimoto thyroiditis, 36 had non-Hashimoto lymphocytic thyroiditis, and 5 had diffuse hyperplasia. The presence of ≥ 2 US and CT features of DTD, which was classified as suspicious for DTD or DTD, had the largest area under the receiver operating characteristic curve (0.866 and 0.893, respectively), with sensitivity and specificity of 71.9 and 91.2% in US, and 84.4 and 84.1% in CT, respectively. However, there was no statistically significant difference between readers' experience and their diagnostic performance.

Conclusion US and CT imaging may be helpful for detecting incidental DTD.

Keywords Thyroid · Diffuse thyroid disease · Ultrasonography · Computed tomography · Multicenter

✉ Dong Wook Kim
dwultra@nate.com

Yoo Jin Lee
netisss@hanmail.net

Hye Shin Ahn
ach0224@gmail.com

Hye Jin Baek
sartre81@gmail.com

Ji Hwa Ryu
rjhrad@empal.com

Taewoo Kang
taewoo.d.kang@gmail.com

[1] Department of Radiology, Busan Paik Hospital, Inje University College of Medicine, 75, Bokji-ro, Busanjin-gu, Busan 47392, South Korea

[2] Department of Radiology, Chung-Ang University Hospital, Chung-Ang University College of Medicine, Seoul 06973, South Korea

[3] Department of Radiology, Gyeongsang National University School of Medicine and Gyeongsang National University Changwon Hospital, Changwon 51476, South Korea

[4] Department of Radiology, Haeundae Paik Hospital, Inje University College of Medicine, Busan 48108, South Korea

[5] Department of Surgery (Busan Cancer Center), Pusan National University Hospital, Pusan National University College of Medicine, Busan 49241, South Korea

Introduction

Thyroid disease can be classified as nodular or diffuse [1]. Nodular thyroid disease is a localized form, whereas diffuse thyroid disease (DTD) is characterized by involvement of the entire thyroid gland, and is a main cause of thyroid dysfunction [2]. Currently, thyroid antibody and thyroid function tests are used for detecting and managing thyroid dysfunction [2, 3]. In clinical practice, however, a diagnostic limitation exists in patients with subclinical hypothyroidism or asymptomatic thyroid dysfunction [2, 3]. Early detection of subclinical DTD can be helpful for appropriate management of thyroid dysfunction [3]. Recently, several radiological approaches for detecting incidental DTD have been proposed using ultrasonography (US) or computed tomography (CT) [4–7]. According to these studies, US or CT was helpful for detecting DTD, although the findings were from single-center investigations. To the best of our knowledge, however, no multicenter study investigating the radiological detection of incidental DTD has been published. In particular, no previous study comparing the diagnostic performance of US and CT using static images has been conducted. Accordingly, the purpose of the present study was to compare the diagnostic performance of CT and US for the detection of incidental DTD using static US and CT images, multicenter data, and histopathological results as a reference standard.

Materials and methods

Patients

The present retrospective analysis was based on patient data collected from four university-affiliated hospitals. The institutional review boards of all participating institutions approved the study. Given the retrospective nature of the study and the use of anonymized data, requirements for informed consent were waived. Between July and December 2016, patients who met the following criteria at each institution were included: underwent thyroid surgery for the treatment of nodular thyroid lesions; received at least one preoperative thyroid US examination; and underwent preoperative neck CT for tumor staging. Ultimately, a total of 177 patients [144 women, 33 men; mean age 45.5 ± 13.0 years (range 15–79 years)] were included in this study.

Thyroid US and retrospective image analysis

All US examinations were performed using one of two high-resolution US systems: iU 22 (Philips Medical Systems, Bothell, WA, USA); or Aplio SSA-770A (Toshiba Medical Systems, Tokyo, Japan). A 5–12 or 8 MHz linear-array transducer was used. Faculty radiologists specializing in head-and-neck imaging or supervised board-certified radiologists participating in head-and-neck radiology fellowship training performed all US examinations.

Four board-certified radiologists (with 5, 6, 10, and 15 years' experience with thyroid US examination after obtaining board certification) performed retrospective analyses of US images using a picture-archiving and communication system. Each radiologist analyzed the US findings from study patients at each institution, and was blinded to the CT diagnosis, clinico-serological information, and patient medication history for DTD. Based on the literature [4, 7], the following features were investigated in the static US images: echogenicity (iso-, low, or high); echotexture (fine, coarse, or micronodulative); anteroposterior diameter of the thyroid gland [normal (1–2 cm), increased (> 2 cm), or decreased (< 1 cm)]; glandular margin (smooth or lobulated); and vascularity (normal, decreased, or increased).

Based on the US features, the enrolled patients were classified into four categories, as follows: those without abnormal US features were classified as no DTD; those with ≥ 1 abnormal US feature were classified as indeterminate; those with ≥ 2 abnormal US features were classified as suspicious for DTD; and those with ≥ 3 abnormal US features were classified as DTD.

Neck CT and retrospective image analysis

For tumor staging before thyroid surgery, neck CT (slice thickness, 3 mm; reconstruction increment, 3 mm) was performed using contrast medium (2–3 mL/s and 40 s delay time) and a 64-channel multi-detector CT scanner (Brilliance 64; Philips Medical Systems, Cleveland, OH, USA and Discovery CT 750HD SP 64; GE Medical Systems, Milwaukee, WI, USA) or a 128-channel multi-detector CT scanner (LightSpeed; GE Medical Systems and SOMATOM Definition AS + ; Siemens Healthcare, Forchheim, Germany). Non-enhanced axial, contrast-enhanced axial, and contrast-enhanced coronal reformatted CT images were acquired in all cases.

Four board-certified radiologists (with 4, 6, 10, and 15 years' experience with neck CT interpretation after obtaining board certification) performed the retrospective analyses of CT images using a picture-archiving and communication system. Each radiologist analyzed the CT findings from study patients at each institution, and was blinded to the US diagnosis, clinico-serological information, and patient medication history for DTD. Based on the literature [5, 6], the CT features of the thyroid gland were retrospectively analyzed in terms of the degree and pattern of parenchymal attenuation in the non-enhanced CT images, glandular size and margin in the contrast-enhanced CT images, and degree and pattern of parenchymal

enhancement in the contrast-enhanced CT images as follows: degree of parenchymal attenuation was divided into iso-, low, or high; patterns of both glandular attenuation and enhancement were categorized into homogeneous or inhomogeneous; anteroposterior diameters of both lobes of the main thyroid were measured, averaged, and classified into 1 of 3 categories: 1–2 cm (normal), < 1, or > 2 cm; margin of the thyroid was classified as smooth or lobulated; degree of parenchymal enhancement was classified as normal, decreased, or increased.

Based on the CT features, the enrolled patients were classified into four categories, as follows: those without abnormal CT features were classified as no DTD; those with ≥ 1 abnormal CT feature were classified as indeterminate; those with ≥ 2 abnormal CT features were classified as suspicious for DTD; and those with ≥ 3 abnormal CT features were classified as DTD.

Histopathology

Histopathological findings from the thyroid gland were classified by board-certified pathologists from each of the affiliated hospitals. The pathologist at each hospital retrospectively reviewed pathological slides with blinded to the US and CT findings. Hashimoto thyroiditis exhibited progressive loss of thyroid follicular cells with replacement by lymphocytes and formation of germinal centers associated with fibrosis. Non-Hashimoto lymphocytic thyroiditis exhibited diffuse infiltration of the thyroid gland with lymphocytes and other inflammatory cells but none of the typical histopathological features of Hashimoto thyroiditis such as oxyphilic metaplasia, follicular atrophy, or follicular disruption. Diffuse hyperplasia exhibited diffuse hypertrophy and hyperplasia of follicular cells with retention of the lobular architecture and no definite nodule formation. The thyroid gland was considered to be normal thyroid parenchyma (NTP) when there was no visual evidence of coexisting DTD.

Statistical analysis

To evaluate the differences in CT and US features between DTD and NTP, independent *t* tests were used for continuous variables and the Pearson's χ^2 test or, for small cell values, Fisher's exact test for categorical variables. Receiver operating characteristic (ROC) curve analysis was applied to obtain the optimal cut-off value of CT and US classifications for detecting DTD. A cut-off value for each variable was determined by maximizing the sum of the sensitivity and specificity. In addition, ROC curve analysis was constructed to evaluate the diagnostic performance for the best predictor of DTD at each institution. The area under the ROC curve (AUC) was compared using the method described by DeLong et al. [8]. Kendall's tau coefficient was also performed to evaluate the linear correlation between reader experience and diagnostic performance, after converting categorical variables of these two factors. The McNemar test was used to compare results between the two imaging modalities. All statistical analyses were performed using SPSS version 24.0 (IBM Corporation, Armonk, NY, USA) and MedCalc version 14.10 (MedCalc, Belgium); $p < 0.05$ was considered to be statistically significant.

Results

Patient characteristics

Based on the histopathological results from 177 patients, 64 (mean age of 45.3 years with an age range of 21–75 years; male:female = 8:56) were subsequently classified as DTD and 113 (mean age 45.6 years [range 15–79 years]; 25 male, 88 female) as NTP. However, there was no significant difference in patient age ($p = 0.879$) between the DTD and NTP groups. The types of thyroid surgery were divided into hemithyroidectomy [100/177 (56.5%)], total thyroidectomy [74/177 (41.8%)], and subtotal thyroidectomy [(3/177 (1.7%)]. After thyroid surgery, papillary thyroid carcinoma [163/177 (92.1%)], follicular thyroid carcinoma [2/177 (1.1%)], follicular adenoma [4/177 (2.3%)], and nodular hyperplasia [8/177 (4.5%)] were identified. On histopathological analysis, NTP [113/177 (63.8%)], Hashimoto thyroiditis [23/117 (13.0%)], non-Hashimoto lymphocytic thyroiditis [36/177 (20.3%)], and diffuse hyperplasia [5/177 (2.8%)] were found. Of the 165 patients with thyroid cancer, NTP [106/165 (64.2%)], Hashimoto thyroiditis [22/165 (13.3%)], non-Hashimoto lymphocytic thyroiditis [33/165 (20%)], and diffuse hyperplasia [4/165 (2.4%)].

Analyses of US and CT features

Comparisons of individual US features between NTP and DTD are summarized in Table 1. The echogenicity, echotexture, glandular size and margin, vascularity, and US classification demonstrated a significant difference between NTP and DTD ($p < 0.05$). US classification in the enrolled patients included no DTD [75/177 (42.4%)], indeterminate [46/177 (26.0%)], suspicious for DTD [28/177 (15.8%)], and DTD [28/177 (15.8%)]. In the comparison of US classification and histopathological results, the 75 cases assigned to the no DTD category included Hashimoto thyroiditis ($n = 2$), non-Hashimoto lymphocytic thyroiditis ($n = 4$), diffuse hyperplasia ($n = 0$), and NTP ($n = 69$). The 46 cases assigned to the indeterminate category included Hashimoto

Table 1 Comparison of ultrasonography features of normal thyroid parenchyma and diffuse thyroid disease in 177 patients

US features	Normal thyroid parenchyma ($n=113$)	DTD ($n=64$)	*P* value
Echogenicity			<0.0001
Iso-	106 (93.8)	34 (53.1)	
Low	7 (6.2)	30 (46.9)	
High	0	0	
Echotexture			<0.0001
Fine	79 (69.9)	9 (14.1)	
Coarse	34 (30.1)	51 (79.7)	
Micronodulative	0	4 (6.3)	
Size of thyroid gland			0.001
Normal	110 (97.3)	54 (84.4)	
Increased	2 (1.8)	10 (15.6)	
Decreased	1 (0.9)	0	
Margin of thyroid gland			<0.0001
Smooth	108 (95.6)	43 (67.2)	
Lobulated	5 (4.4)	21 (32.8)	
Vascularity			<0.0001
Normal	105 (92.9)	35 (54.7)	
Decreased	1 (0.9)	2 (3.1)	
Increased	7 (6.2)	27 (42.2)	
Thyroid inferno	0	0	
US classification			<0.0001
No DTD	69 (61.1)	6 (9.4)	
Indeterminate	34 (30.1)	12 (18.8)	
Suspicious for DTD	7 (6.2)	21 (32.8)	
DTD	3 (2.7)	25 (39.1)	

Data presented in parentheses are percentage of each item

US ultrasonography, *DTD* diffuse thyroid disease

Table 2 Comparison of computed tomography features of normal thyroid parenchyma and diffuse thyroid disease in 177 patients

CT features	Normal thyroid parenchyma ($n=113$)	DTD ($n=64$)	*P* value
Degree of attenuation			<0.0001
Iso-	96 (85)	27 (42.2)	
Low	16 (14.2)	37 (57.8)	
High	1 (0.9)	0	
Pattern of attenuation			<0.0001
Homogeneous	83 (73.5)	15 (23.4)	
Inhomogeneous	30 (26.5)	49 (76.6)	
Size of thyroid gland			<0.0001
Normal	106 (93.6)	40 (62.5)	
Increased	7 (6.2)	24 (37.5)	
Decreased	0	0	
Margin of thyroid gland			<0.0001
Smooth	108 (95.6)	46 (71.9)	
Lobulated	5 (4.4)	18 (28.1)	
Degree of enhancement			<0.0001
Normal	109 (96.5)	46 (71.9)	
Decreased	4 (3.5)	17 (26.6)	
Increased	0	1 (1.6)	
Pattern of enhancement			<0.0001
Homogeneous	104 (92)	16 (25)	
Inhomogeneous	9 (8)	48 (75)	
CT classification			<0.0001
No DTD	67 (59.3)	4 (6.3)	
Indeterminate	28 (24.8)	6 (9.4)	
Suspicious for DTD	14 (12.4)	17 (26.6)	
DTD	4 (3.5)	37 (57.8)	

Data presented in parentheses are percentage of each item

CT computed tomography, *DTD* diffuse thyroid disease

thyroiditis ($n=3$), non-Hashimoto lymphocytic thyroiditis ($n=9$), diffuse hyperplasia ($n=0$), and NTP ($n=34$). The 28 cases assigned to the suspicious for DTD category included Hashimoto thyroiditis ($n=8$), non-Hashimoto lymphocytic thyroiditis ($n=12$), diffuse hyperplasia ($n=1$), and NTP ($n=7$). The 28 cases assigned to the DTD category included Hashimoto thyroiditis ($n=10$), non-Hashimoto lymphocytic thyroiditis ($n=11$), diffuse hyperplasia ($n=4$), and NTP ($n=3$).

Comparisons of individual CT features between NTP and DTD are summarized in Table 2. The degree and pattern of parenchymal attenuation, glandular size, degree and pattern of parenchymal enhancement, thyroid margin, and CT classification demonstrated a significant difference between NTP and DTD ($p<0.0001$). CT classifications in the enrolled patients included no DTD [71/177 (40.1%)], indeterminate [34/177 (19.2%)], suspicious for DTD [31/177 (17.5%)], and DTD [41/177 (23.2%)]. In the comparison of CT classification and histopathologic results, the 71 cases assigned to the no DTD category included Hashimoto thyroiditis ($n=2$), non-Hashimoto lymphocytic thyroiditis ($n=2$), diffuse hyperplasia ($n=0$), and NTP ($n=67$). The 34 cases assigned to the indeterminate category included Hashimoto thyroiditis ($n=1$), non-Hashimoto lymphocytic thyroiditis ($n=5$), diffuse hyperplasia ($n=0$), and NTP ($n=28$). The 31 cases assigned to the suspicious for DTD category included Hashimoto thyroiditis ($n=6$), non-Hashimoto lymphocytic thyroiditis ($n=11$), diffuse hyperplasia ($n=0$), and NTP ($n=14$). The 41 cases assigned to the DTD category included Hashimoto thyroiditis ($n=14$), non-Hashimoto lymphocytic thyroiditis ($n=18$), diffuse hyperplasia ($n=5$), and NTP ($n=4$) (Figs. 1, 2).

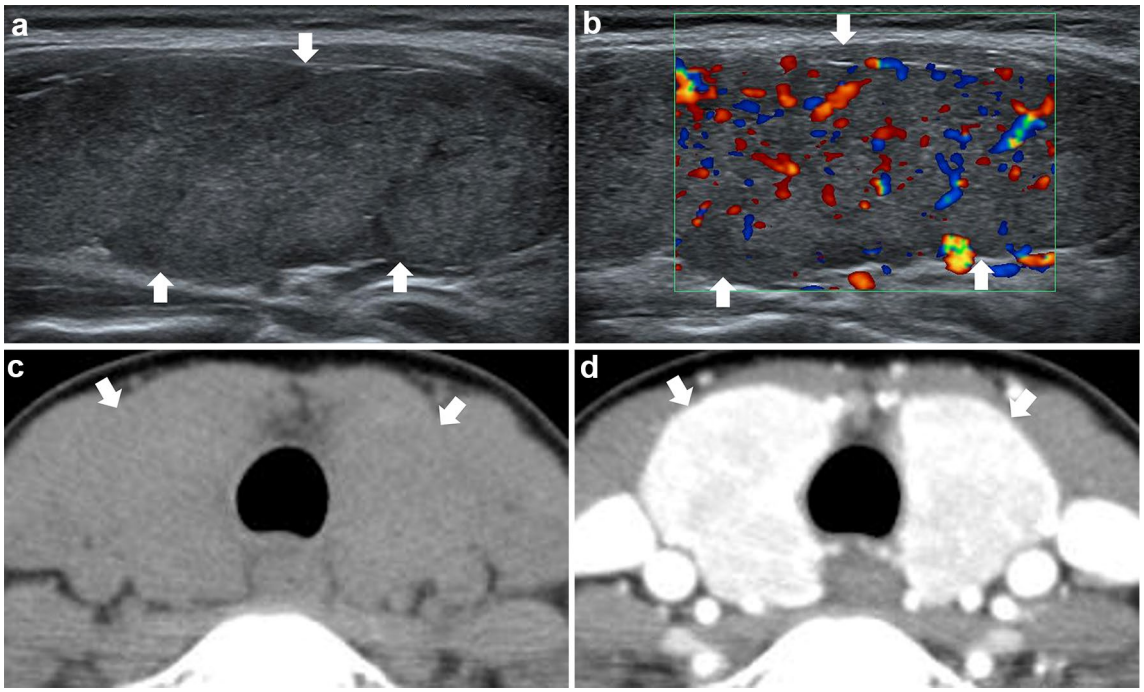

Fig. 1 Ultrasonography (US) and computed tomography (CT) images from a 32-year-old woman diagnosed with Hashimoto thyroiditis on histopathological examination, which was determined to be diffuse thyroid disease (DTD) category on both US and CT (papillary thyroid carcinoma in the right thyroid). In the longitudinal gray-scale (**a**) and color Doppler (**b**) sonograms, the thyroid gland (arrows) exhibits hypoechogenicity, coarse echotexture, anteroposterior diameter > 2 cm, lobulated margin, and increased vascularity. In the non-enhanced (**c**) and contrast-enhanced (**d**) axial CT images, the thyroid gland (arrows) shows low, but homogeneous attenuation, anteroposterior diameter > 2 cm, smooth margin, and normal but inhomogeneous enhancement

Diagnostic performance of US and CT classifications

In ROC curve analyses, the classification of ≥ 2 features of DTD was to be the best predictor of DTD on both of US and CT diagnoses. The classification of ≥ 2 features of DTD on both US and CT demonstrated the largest AUC as follows: US diagnosis [AUC 0.866 (95% CI 0.807 to 0.912)], with a sensitivity of 71.9% and a specificity of 91.2%, and CT diagnosis [AUC 0.893 (95% CI 0.838 to 0.935)], with a sensitivity of 84.4% and a specificity of 84.1% (Fig. 3). There were no statistically significant differences in all values between US and CT diagnoses [$p > 0.05$ (McNemar test)].

Analyses of US and CT diagnoses according to individual institution

US classifications for detecting DTD at each institution are summarized in Table 3. According to the comparative results, US classification demonstrated a significant difference between DTD and NTP in all participating institutions ($p < 0.05$), and revealed a positive linear correlation between the detection of DTD and US classification ($p < 0.0001$). In all institutions, the prevalence of the indeterminate category was 39.0% (46/177), ranging from 17.1% (7/41) to 37.5% (15/40). When the indeterminate category cases were excluded, the prevalence of false-negative diagnosis for detecting DTD was low (mean 4.6% [6/131]; range 0% [0/25] to 8.8% [3/34]) and the prevalence of false-positive diagnosis was variable [mean 7.6% (10/131); range 2.9% (1/34) to 14.3% (5/35)].

CT classifications for detecting DTD at each institution are summarized in Table 4. According to the comparative results, CT classification demonstrated a significant difference between DTD and NTP in all participating institutions ($p < 0.0001$), and revealed a positive linear correlation between the detection of DTD and CT classification ($p < 0.0001$). In all institutions, the prevalence of the indeterminate category was 19.2% (34/177), ranging from 12.2% (5/41) to 35.6% (16/45). When the indeterminate category cases were excluded, the prevalence of false-negative diagnosis for detecting DTD was low [mean, 2.8% (4/143); range 0–6.9%) and the prevalence of false-positive diagnosis was variable [mean 12.6% (18/143); range 0% (0/29)–21.1% (8/35)].

The diagnostic performance of each reader at the participating institutions for detecting DTD on US and CT was also analyzed and is summarized in Table 5. Among the four institutions, three demonstrated higher sensitivity in CT and higher specificity in US, whereas one demonstrated similar sensitivity and specificity between CT

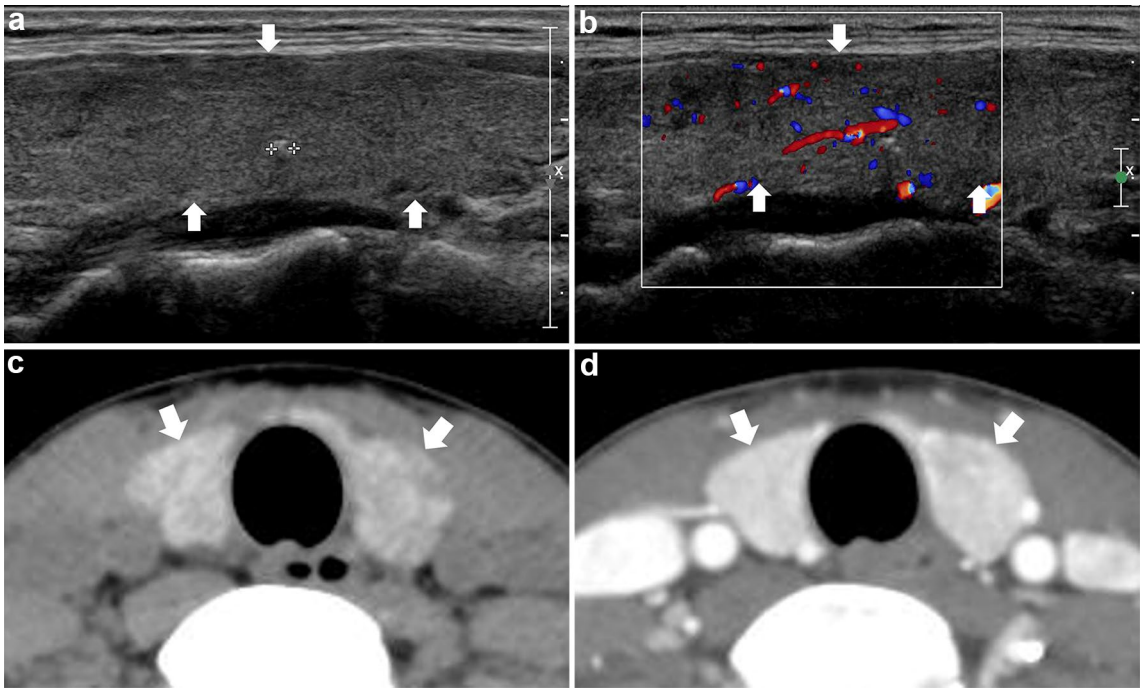

Fig. 2 Ultrasonography (US) and computed tomography (CT) images from a 36-year-old woman diagnosed with non-Hashimoto lymphocytic thyroiditis on histopathological examination, which was determined to be no diffuse thyroid disease (DTD) category on US, but DTD category on CT (papillary thyroid carcinoma in the left thyroid). In the longitudinal gray-scale (**a**) and color Doppler (**b**) sonograms, the thyroid gland (arrows) demonstrates isoechogenicity, fine echotexture, anteroposterior diameter of 1–2 cm, smooth margin, and normal vascularity. In the non-enhanced (**c**) and contrast-enhanced (**d**) axial CT images, the thyroid gland (arrows) exhibits low and inhomogeneous attenuation, anteroposterior diameter of 1–2 cm, smooth margin, and decreased but homogeneous enhancement

and US. In all institutions, positive predictive values were higher for US than CT. However, negative predictive values were higher for CT than US, except for one institution with similar positive and negative predictive values. There was no statistically significant difference between readers' experience and their diagnostic performance ($\tau = 0.21$; $p = 0.764$).

Discussion

In clinical practice, subclinical hypothyroidism or asymptomatic thyroid dysfunction is a diagnostic dilemma [2, 3]. To our knowledge, routine thyroid antibodies and function tests are not performed to detect subclinical hypothyroidism or asymptomatic thyroid dysfunction because of cost issues. Recent studies have reported that US or CT may be feasible methods for detecting subclinical DTD [4, 5, 7]. In the literature, US features of DTD include low echogenicity, coarse or micronodulative echotexture, lobulated margin, and increased parenchymal vascularity [4, 7], whereas CT features of DTD include low and inhomogeneous attenuation, increased gland size, and increased and inhomogeneous enhancement [5, 6]. In the present study, we used US and CT images to differentiate DTD from NTP, and similar findings were observed. Moreover, there was no US or CT feature of DTD demonstrating both high sensitivity and specificity. Similar to previous studies, normal gland size and smooth margin were commonly observed in DTD on both US and CT [4–7].

In the present study, ≥ 2 features of DTD demonstrated the highest diagnostic value in both US and CT, similar to

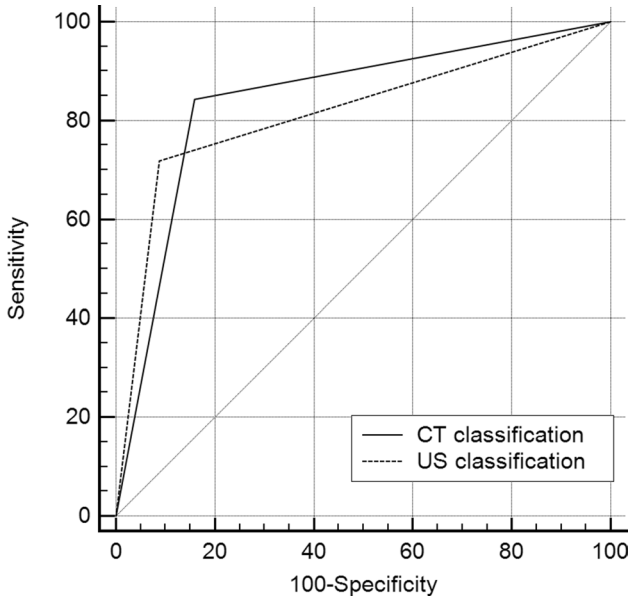

Fig. 3 Comparisons of receiver operating characteristic (ROC) curves representing the diagnostic performance of ultrasonography and computed tomography classifications as the best independent predictor for differentiating diffuse thyroid disease from normal thyroid parenchyma. Diagonal line = 50% of the area under the ROC curve (AUC), and also refers to a hypothetical marker that has no discriminatory power for differentiating diffuse thyroid disease from normal thyroid parenchyma: US [AUC 0.866 (95% CI 0.807–0.912)] and CT [AUC 0.893 (95% CI 0.838–0.935)]

Table 3 Comparison of ultrasonography classification for detecting diffuse thyroid disease according to institution

Institution	US classification	Normal thyroid parenchyma	DTD	*P* value
A ($n=45$)				0.001
	No DTD	12 (52.2)	1 (4.5)	
	Indeterminate	6 (26.1)	4 (18.2)	
	Suspicious for DTD	3 (13)	8 (36.4)	
	DTD	2 (8.7)	9 (40.9)	
B ($n=51$)				<0.0001
	No DTD	21 (60)	2 (12.5)	
	Indeterminate	12 (34.3)	2 (12.5)	
	Suspicious for DTD	1 (2.9)	7 (43.8)	
	DTD	1 (2.9)	5 (31.3)	
C ($n=40$)				<0.0001
	No DTD	14 (56)	0	
	Indeterminate	9 (36)	6 (40)	
	Suspicious for DTD	2 (8)	5 (33.3)	
	DTD	0	4 (26.7)	
D ($n=41$)				<0.0001
	No DTD	22 (73.3)	3 (27.3)	
	Indeterminate	7 (23.3)	0	
	Suspicious for DTD	1 (3.3)	1 (9.1)	
	DTD	0	7 (63.6)	

Data presented in parentheses are percentage of each item

CT computed tomography, *DTD* diffuse thyroid disease

previous US studies reported in the literature [4, 5]. However, previous CT studies have reported different results [5, 6]. One reported the highest diagnostic values in ≥ 2 CT features of DTD; the other revealed the highest diagnostic values in ≥ 3 CT features. In the present study, ≥ 2 CT features of DTD demonstrated the highest diagnostic value. The reason for this difference is unclear. To clarify the cut-off number of abnormal CT features for detecting DTD, further study may be required.

To our knowledge, only one comparative study of US and CT for detecting DTD has been reported [5]. This single-center study used real-time US, and reported similar diagnostic values between US and CT. However, no comparative study involving retrospective analyses of US images is available. In the present study, we used static images for both US and CT diagnoses. There was no statistically significant difference in diagnostic value between US and CT, although positive and negative predictive values were better in US and CT, retrospectively. The reason for this is unclear; however, a possible explanation is that our results correspond with results from a previous study in that real-time US was superior to static US for detecting incidental DTD [7], and that real-time US and CT demonstrated similar diagnostic values [5]. In all participating institutions, the number of static US images was significantly lower than the number of CT images. Furthermore, the prevalence of the indeterminate category was higher for US [26.0% (46/177)] than for CT [19.2% (34/177)].

Table 4 Comparison of computed tomography classification for detecting diffuse thyroid disease according to institution

Institution	CT classification	Normal thyroid parenchyma	DTD	*P* value
A ($n=45$)				<0.0001
	No DTD	10 (43.5)	2 (9.1)	
	Indeterminate	13 (56.5)	3 (13.6)	
	Suspicious for DTD	0	4 (18.2)	
	DTD	0	13 (59.1)	
B ($n=51$)				<0.0001
	No DTD	22 (62.9)	0	
	Indeterminate	6 (17.1)	2 (12.5)	
	Suspicious for DTD	6 (17.1)	5 (31.3)	
	DTD	1 (2.9)	9 (56.3)	
C ($n=40$)				<0.0001
	No DTD	12 (48)	0	
	Indeterminate	5 (20)	0	
	Suspicious for DTD	6 (24)	8 (53.3)	
	DTD	2 (8)	7 (46.7)	
D ($n=41$)				<0.0001
	No DTD	23 (76.7)	2 (18.2)	
	Indeterminate	4 (13.3)	1 (9.1)	
	Suspicious for DTD	2 (6.7)	0	
	DTD	1 (3.3)	8 (72.7)	

Data presented in parentheses are percentage of each item

CT computed tomography, *DTD* diffuse thyroid disease

Thus, we believe that CT is a useful tool for differentiating DTD from NTP.

This study had several limitations, the first of which was the small sample size at each institution. This may have compromised statistical interpretation of the results. Second, all study patients underwent thyroid surgery, which may have resulted in unavoidable selection bias. Third, awareness of CT features of DTD or experience with CT diagnosis of DTD was not evaluated; only radiologist experience with neck CT interpretation was investigated. Fourth, clinico-serological data regarding thyroid function were not included. Fifth, retrospective analyses of US and CT images were performed at all participating institutions. However, the number of US images for the thyroid gland was variable according to each case or institution; the effect of this diversity was not evaluated. Furthermore, the number of US images for the thyroid gland was smaller in comparison with CT images. Finally, there were differences among the affiliated hospitals with regard to US and CT modalities, and CT protocols.

In conclusion, both US and CT demonstrated similar diagnostic value for differentiating DTD from NTP. Also, US and CT images may be helpful for detecting incidental DTD, especially when ≥ 2 features of DTD are apparent.

Table 5 Diagnostic index of US and CT for detecting diffuse thyroid disease according to institution (when "2 or more" US and "2 or more" CT classifications were selected)

Institutions		A_z value[a]	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	*P* value
A	US	0.778 (0.629, 0.888)	77.3	78.3	77.3	78.3	<0.0001
	CT	0.672 (0.516, 0.804)	90.9	43.5	60.6	83.3	0.005
B	US	0.846 (0.718, 0.932)	75	94.3	85.7	89.2	<0.0001
	CT	0.838 (0.708, 0.926)	87.5	80	66.7	93.3	<0.0001
C	US	0.760 (0.599, 0.881)	60	92	81.8	79.3	0.0003
	CT	0.840 (0.690, 0.936)	100	68	65.2	100	<0.0001
D	US	0.847 (0.700, 0.940)	72.7	96.7	88.9	90.6	<0.0001
	CT	0.814 (0.661, 0.918)	72.7	90	72.7	90	<0.0001
Total	US	0.866 (0.807, 0.912)	71.9	91.2	82.1	85.1	<0.0001
	CT	0.893 (0.838, 0.935)	84.4	84.1	75	90.5	<0.0001

[a] A_z means the largest area under the Receiver operating characteristic curve

US ultrasonography, *CT* computed tomography, *PPV* positive predictive value, *NPV* negative predictive value

Acknowledgment This word was supported by the 2017 Inje University research Grant.

Authors' contributions Concept and design: DWK. Acquisition of data: All authors. Literature review: All authors. Analysis and interpretation of data: DWK, HJB. Manuscript writing: DWK. Refinement of manuscript: All authors. Review of final manuscript: DWK. Final approval: DWK.

Funding None.

Compliance with ethical standards

Conflict of interest All authors declare that there are no competing interests.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent The need for informed consent was waived owing to the retrospective nature of the study.

References

1. Frates MC, Benson CB, Charboneau JW, Cibas ES, Clark OH, Coleman BG, Cronan JJ, Doubilet PM, Evans DB, Goellner JR, Hay ID, Hertzberg BS, Intenzo CM, Jeffrey RB, Langer JE, Larsen PR, Mandel SJ, Middleton WD, Reading CC, Sherman SI, Tessler FN, Society of Radiologists in Ultrasound (2005) Management of thyroid nodules detected at US: society of radiologists in ultrasound consensus conference statement. Radiology 237:794–800
2. Pearce EN, Farwell AP, Braverman LE (2003) Thyroiditis. N Engl J Med 348:2646–2655
3. Rosario PW, Bessa B, Valadao MM, Purisch S (2009) Natural history of mild subclinical hypothyroidism: prognostic value of ultrasound. Thyroid 19:9–12
4. Kim DW, Eun CK, In HS, Kim MH, Jung SJ, Bae SK (2010) Sonographic differentiation of asymptomatic diffuse thyroid disease from normal thyroid: a prospective study. AJNR Am J Neuroradiol 31:1956–1960
5. Kim DW, Jung SJ, Ha TK, Park HK, Kang T (2014) Comparative study of ultrasound and computed tomography for incidentally detecting diffuse thyroid disease. Ultrasound Med Biol 40:1778–1784
6. Rho MH, Kim DW (2014) Computed tomography features of incidentally detected diffuse thyroid disease. Int J Endocrinol 2014:921934
7. Kim DW (2015) A comparative study of real-time and static ultrasonography diagnoses for the incidental detection of diffuse thyroid disease. Endocr Pract 21:910–916
8. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845